图书情报工作
第63卷 第6期 2019年3月

专利文献中的睡美人现象分析与研究

■ 李贺　袁翠敏　解梦凡
吉林大学管理学院　长春 130022

摘要：［目的/意义］“睡美人”文献是对科学论文中存在的迟滞认可现象的描述，而延迟发现与延迟关注的现象也同样存在于技术文献中。在梳理文献中的睡美人、专利沉睡现象及专利引文分析的相关研究后，将此概念引入到专利信息分析中，揭示专利文献中存在的睡美人现象。［方法/过程］以美国专利商标局和美国国家经济研究局发布的专利及其引证信息为基础，使用睡美人文献经典识别方法识别出睡美人专利，对其进行特征分布分析，并选取典型案例进行研究。［结果/结论］结果证明专利文献中也存在睡美人现象，且拥有专利文献特有的特征，为后续睡美人专利的识别与唤醒奠定基础，进而为及早发现并利用此类有价值的专利文献提供解决方案，促进知识流动和技术迭代，提高科研效率，加速科学发现。

关键词：延迟认可　睡美人文献　专利引文分析　睡美人专利
分类号：G250
DOI：10.13266/j.issn.0252-3116.2019.06.009

1 引言

科学论文中存在延迟承认的“睡美人”文献，专利文献中也存在类似现象的“沉睡专利”和“潜水艇专利”。这些专利文献起初并未获得广泛关注，随着技术的发展演进和专利战略的实施，受专利权保护的技术成为某一领域的重要力量。由此可见，延迟承认或延迟关注、延迟发现的现象，不仅存在于以论文为代表的科学文献中，也存在于以专利文献为主的技术文献中[1]。睡美人文献的识别与研究基于引文分析。与论文的参考文献结构与引用方式类似，专利文献也包含了引用与被引用关系。专利发明人要依照法律规定披露作为本专利研究基础的已有专利，申请发明专利方面还要对作为技术基础的工艺水平进行完整的描述[2]。不同的是，科学论文的引用由文章作者负责标注，而专利文献的引用不仅包含专利发明人的标注，专利审查人、专利代理人乃至社会公众也要根据实际引证情况与专利本身的特征，遵照专利法的相关要求做相应的标注。专利引证关系作为技术的链接，在专利分析计量领域具有重要价值，可用于揭示技术流动及扩散、绘制知识结构地图、用于追踪科学技术发展等方面[3]。

以往对沉睡专利的研究着眼于市场竞争和专利战略布局，本文则尝试将睡美人文献的识别和分析方法与专利引文分析相结合，回归专利的文献属性，从引文视角对沉睡专利进行研究，以揭示专利文献中存在的睡美人现象。

2 文献综述

2.1 文献中的“睡美人”现象及含义

在科学研究中，科学出版物的价值通常会在文献出版之后很快获得认可，其引用量逐年增加，并因主题领域、期刊类型以及文献内容的不同，引用量在文献发表2-6年之后达到峰值，期间成为领域研究热点，而随着时间的推移和学科发展，文献引用量逐年减少直至不再被引用[4]。这类典型文献的引用曲线基本符合正态分布[5]。然而，还有一部分文献在发表之初，其重要性和价值并没有得到认可，在经历了数年之后，才开始获得关注，文献被大量引用。学者对此类现象进行了研究，分别将其命名为抵制发现[6]（resisted discoveries）、早熟发现[7]（premature discoveries）、迟滞认可[8]（delayed recognition）等，学者把产生这种现象的原因总结为研究成果过于超前，不能与当时的知识水平保持一致，或不能被当时的技术所证实，此外，这种现象也与研究主题、科学社会或科学家对相关研究所持态

作者简介：李贺（ORCID：0000-0001-8847-3619），教授，博士生导师；袁翠敏（ORCID：0000-0001-8415-838X），博士研究生，通讯作者，E-mail：yuancuimin@126.com；解梦凡（ORCID：0000-0002-0670-8160），博士研究生。
收稿日期：2018-05-31　修回日期：2018-09-07　本文起止页码：64-74　本文责任编辑：王传清

度有关等[6]。A. F. J. V. Raan 于 2004 年借鉴童话故事,将该现象定义为“睡美人文献”,并获得了广泛认可:科学中的睡美人是指出版物发表之后很久没有受到关注(沉睡),而后几乎是突然之间吸引了大量的关注(由王子唤醒)[9],典型引文曲线见图 1。最经典的睡美人文献是孟德尔于 1865 年发表的论文《植物杂交实验》,这篇文章在发表了 34 年之后才被研究人员重新验证[10]。睡美人文献所占比例较小,但其价值却是巨大的。

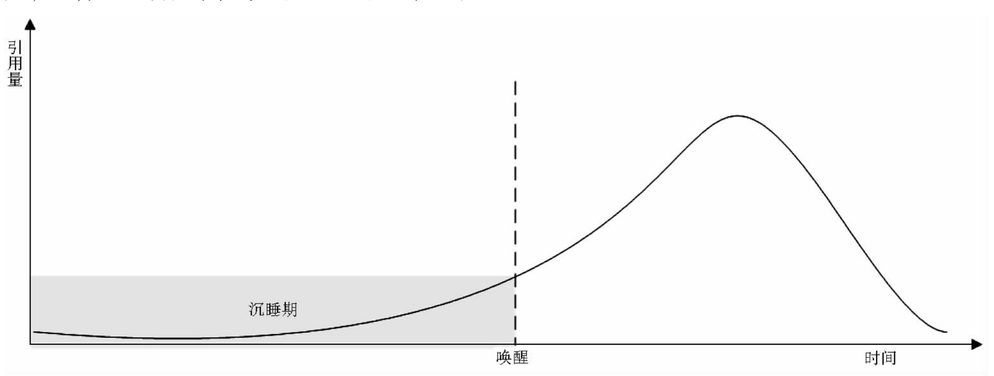


图 1　睡美人文献典型引用曲线

目前,学者对睡美人文献的研究已初具规模,在睡美人的概念、特征描述、识别方法、形成原因、唤醒机制及案例研究等方面都取得了较大进展。在概念方面,自 A. F. J. V. Raan 提出经典定义之后,学者又对其不断进行完善。如 2012 年,我国学者李江、叶鹰根据引文曲线的特征,结合文献引用所呈现的“昙花一现”的现象,进一步丰富了睡美人的概念内涵,提出了“全要素”睡美人的概念[11]。睡美人文献的识别方法研究较为丰富,主要可以分为定量识别方法和曲线拟合识别方法。其中,定量识别方法又可以分为人为设定参数的定量方法(以 A. F. J. V. Raan 提出的三指标识别方法[9]和 R. Costas 提出的四分位法[12]及其改良方法等为主)以及无参数客观指标定量方法(如“睡美人指数”法等[13]);引文曲线识别方法,如李江构建的引文曲线分析框架[14]。此外,J. Wang 等针对睡美人文献研究存在的片面性等问题,使用案例分析方法,揭示了睡美人被唤醒的原因及唤醒机制,并从更普遍意义上,提出唤醒策略[15]。杜建、武夷山等也系统研究了睡美人文献与王子文献的识别方法[16],以及睡美人文献的唤醒机制[1]。可见睡美人文献研究已成体系,且形成了相对成熟的理论与方法。

2.2　专利沉睡现象研究

专利文献的沉睡现象是指专利在申请授权之后便不再被使用,或没有对其他专利产生启发作用,更多的是指专利没有被商业化并应用到实际生产中。以一个典型案例来说明专利沉睡现象:中国科学院过程工程研究所曾于 1995 年申请了一项解耦燃烧技术的专利,该技术主要涉及采暖炉燃烧和废气排放,并于 1998 年获得授权,但直到 2017 年冬天该技术成果才得以大面积推广应用。专利申请之初,国内还未意识到氮氧化物的危害,该技术明显超前于市场。经过 20 年的发展该专利才获得重视并被推广应用,专利和技术的价值得以发挥。由此可见,延迟发现不仅存在于科学论文中,技术文献中兼而有之。

国外学者从市场竞争和专利竞赛的角度对沉睡专利开展了大量研究。在企业竞争中,企业为了保证垄断地位,在潜在竞争对手之前尽可能多而全地对新技术申请专利,这种竞争活动可能会导致有些专利既没有使用也没有授权给其他企业的情况出现,这类专利被称之为沉睡专利[17]。在特定市场环境中,企业的先发制人、优先占有专利申请的市场策略可能导致沉睡专利的出现[18]。有研究发现沉睡专利可能更具有创新性,并值得采取措施将其唤醒[19]。但企业强制实施沉睡专利战略的举措,在掣肘竞争对手的同时,也可能在一定程度上损害企业的技术研发动机[20]。我国学者唐家要于 2006 年提出“专利沉睡”的概念[21],此后学者也对专利沉睡的现象、原因及唤醒机制等进行了研究,将专利定义为一种资产,并依据专利应用程度将其分为活跃专利、阻碍专利以及沉睡专利,其中存在质量不高或不符合专利保护相关规范的问题专利[22],见图 2。形成沉睡专利的原因,学者总结为利益差异、制度过度干预、专利产生机制不完善与知识产权保护问题以及专利交易市场信息不对称等问题[23]。在沉睡专利的唤醒上,专利质量评价及专利权有效交易是重

要途径，此外还包括平衡利益关系、完善法律法规及干预制度等。

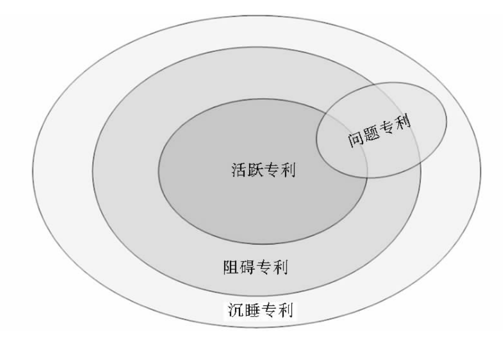


图 2 专利应用程度分类

专利沉睡现象中存在一类特殊专利，即“潜水艇专利”。这类专利是指专利（受专利制度限制，尤指美国专利）在申请后，以各类缘由提出延续申请，延长专利申请时间，以在产业技术成熟之前，尽可能长时间保守发明专利的秘密。而一旦技术变革，产业发展成熟，随即获得专利权，专利权人以此规避他人使用该专利的权利，获取丰厚的专利权转让费用[24]。但此类专利的沉睡，是人为蓄意造成的，是专利战略和专利布局的重要方式[25]，对企业获取高额回报具有重要意义。但在一定程度上是垄断竞争的行为，不利于科学发展和技术演进，存在巨大争议[26]，专利申请审查制度也随之发生改变以遏制相关行为[27]。

以往对专利沉睡现象的研究都是从商和法的视角对专利的利用和转化进行的讨论，其实质是将专利作为一类资产来研究。本文着眼于专利的文献视角，回归专利的文献特征来进行研究。在专利申请授权之后，长时间没有受到关注，未对其他专利产生启发或没有成为技术迭代的基础专利，也是沉睡的表现，主要体现在没有被其他专利文献引用。已有对专利文献的“零被引”和“低被引”的研究是基于某个时间点的引证数据，分层分析专利技术的发明创造水平，研究发现“零被引”专利代表了微型发明创造技术[28]。本文则是根据专利文献引用的历史数据和时间分布特征进行分析。因此，本文拓展了沉睡专利的概念，从专利引证和文献计量学的角度，追本溯源研究延迟关注的沉睡专利，对专利的研究与利用提出新的研究视角。

2.3 专利计量与专利引文分析

早在 20 世纪 40 年代就有学者提出使用计量方法分析研究专利信息[29]，随着专利信息的存储日益规范、可获得性逐步提高，研究人员将研究的焦点放在专利信息本身所具有的属性上，推进专利信息分析思路的建立与发展[30]。20 世纪 80 年代，美国知名专利信息分析公司——美国知识产权咨询公司（CHI Research）将文献计量方法借鉴到专利信息分析领域，拓展了该领域的研究方法与思维方式，1994 年，F. Narin 发表的文章中正式使用计量方法分析专利信息[31]，成为领域先驱。随后，F. Narin 的研究更为广泛和深入，逐步确立了专利计量的研究方法，同时，专利计量也丰富了文献计量学的方法。

与科学论文类似，专利作为重要的科技文献，若长期未被引用或转化，也可能存在被忽视的且极具价值的研究成果或是具有研究潜力的重要内容。低被引专利并不能被简单地归为“垃圾”专利，其价值或许未被发现。专利计量研究主要包括对专利数量、专利引用及专利关联关系 3 个方面的分析与研究[32]，可见专利引文分析的重要程度。通过专利引文分析，可发掘专利技术价值、掌握技术发展轨迹与知识扩散路径[33]，并可通过对引文的时间序列分析测定技术生命周期与成熟度。专利引证指标主要包括引文数量与被引次数，随着专利制度与分析方法的演化，又拓展出专利参考文献与非专利参考文献的数量研究、科学关联度、当前影响指数和技术强度[34]等。专利引证分析在宏观、中观和微观分析角度都具有很好的表征作用[35]，为评价企业发展状态，进而支持企业发展决策起到重要作用。

我国专利申请和授权量逐年增加，但专利的利用率和转化率并不高，在很大程度上制约了经济的健康平稳发展，并造成了研发资源的浪费[36]。本文尝试将文献睡美人的相关理论引入到专利信息分析中，利用专利引证信息识别专利睡美人文献，以探究专利文献中的睡美人现象。

3 研究方法与数据基础

专利计量与文献计量之间的渊源颇深，专利计量发展之初就延用了文献计量学的诸多方法和理论，并且专利文献与科学论文之间在文档结构上存在诸多相似之处。因此，在专利计量中借鉴文献计量领域的新理论、新方法是有迹可循的。睡美人文献是近年来学术界对科学论文中存在迟滞认可现象的总结，是文献计量领域的新概念，将睡美人的概念引入到专利计量领域，有利于丰富专利信息分析尤其是专利引证分析的方法与思路。本文以专利信息为基础，包括专利基本信息与专利引证信息，使用睡美人文献的经典识别方法尝试识别其中可能存在的睡美人专利，结合睡美

李贺，袁翠敏，解梦凡. 专利文献中的睡美人现象分析与研究[J]. 图书情报工作,2019,63(6):64－74.

人文献的相关指标与专利信息本身特征,对识别出的睡美人专利从整体上进行特征分析,选取睡美人专利中存在的典型案例进行剖析,以揭示专利中存在的睡美人现象,为后续睡美人专利的进一步研究奠定基础。

睡美人文献识别方法以 A. F. J. V. Raan 提出的三指标识别方法最为经典,以此方法识别出的睡美人文献特征较为典型,在科学睡美人的相关研究中得到广泛使用。本文将在专利睡美人文献的识别中使用该方法。A. F. J. V. Raan 提出的睡美人文献的三个指标,分别是睡眠深度(depth of sleep)、睡眠长度(length of sleep)以及唤醒强度(awake intensity),其含义如下:

睡眠深度(c_s):是指文献睡眠期间的年均被引量,分为深度睡眠和浅度睡眠。深度睡眠是指文献在睡眠期间年均被引量至多为 1 次,浅度睡眠是指文献在睡眠期间年均被引量在 1－2 次。

睡眠长度(s):是指文献深度睡眠或浅度睡眠所经历的时长。

唤醒强度(c_w):是指唤醒后 4 年的年均引用量。

A. F. J. V. Raan 在对不同参数下的文献数据分析的基础上,得到了睡美人公式如下:

$$N = f\{s, c_s, c_w\} \sim s^{-2.7} \cdot c_s^{+2.5} \cdot c_w^{-6.6} \qquad 公式(1)$$

依据 A. F. J. V. Raan 及相关研究的参数设置,本文将睡眠深度年均引用量设定为 $c_s \leqslant 2$ 次;睡眠长度设定为 $s \geqslant 5$ 年;唤醒强度为唤醒后 4 年的年均引用量设定为 $c_w \geqslant 5$ 次。

本文的数据基础是美国国家经济研究局(National Bureau of Economic Research, NBER)发布的专利及专利引证信息数据库,即专利数据计划(NBER Patent Data Project)。NBER 于 2001 年发布了《专利引证数据文件的教程、洞见及方法论工具》,对数据内容、专利引文分析指标与分析方法和分析案例进行了说明[37]。该数据库涵盖了美国专利商标局(USPTO)自 1976 年至 2006 年的所有发明专利及其引证关系,其中专利数据 485 万余条(样本数据见表 1),主要包含了专利号、专利权人代码、申请年、授权年、国际专利分类号等主要字段;专利引证数据 2 365 万余对(样本数据见表 2),每一条记录表达了一个引证关系中的施引专利和被引专利,并用专利号进行标识,成对标识的方法为前向引证的数据统计与分析提供了便捷的条件。该数据库对于尝试性识别专利睡美人文献来说数据量充足、有效,同时也为睡美人专利的特征分析提供了必要的数据基础。由于美国专利自申请日或优先权日起有 18 个月的保护期,数据库中存储了各个专利的申请年和授权年,识别时选择授权年作为被引专利的公开时间起点,申请年为施引专利的施引时间,以保证睡美人专利识别的准确性与合理性。

表 1　专利数据信息样表示例

序号	专利号	申请年	专利权人代码	授权年	国际专利分类号	主分类号	主要专利类	技术子类
1	3930271	1974	251415	1976	A41D 1900	A41D	2	63
2	3930272	1974	246000	1976	A47D 701	A47D	5	65
3	3930273	1975	10490	1976	A47C 2100	A47C	5	65
4	3930273	1975	10490	1976	A47D 702	A47D	5	65
5	3930274	1974	0	1976	B63B 708	B63B	114	55

表 2　专利引证对样表示例

序号	施引专利的专利号	被引专利的专利号
1	5135226	3930271
2	5592695	3930271
3	5621918	3930271
4	5624296	3930271
5	5675839	3930271

4　专利睡美人文献特征分析与典型案例研究

通过将 A. F. J. V. Raan 的睡美人文献识别方法应用到专利领域,笔者从 2 365 万余对引证数据和 485 万余条专利数据中,经过大批量数据库运算,识别出专利睡美人文献 15 016 件,占比约为 0. 31%。已有研究从普通论文中识别出的睡美人论文比例约为 0. 01% － 0. 1%[1],由此可见,就当前识别方法下,专利文献中的睡美人比例大于科学论文中睡美人文献的比例。NBER 的专利数据 2006 年之后不再更新,但对早期专利个案的前向引证年度分布影响并不大。为了更好地观察所识别出的专利睡美人文献的年度被引情况及曲线特征,笔者在美国专利商标局(USPTO)网站以“Referenced By”字段检索专利的前向引证专利信息,以更新相应的专利信息,弥补 NBER 专利数据的时间差,并对检索的专利结果年度分布情况进行了统计,用以绘制典型案例的引证曲线。

4.1 专利睡美人文献特征分析

4.1.1 专利睡美人文献申请年度分布 通过对 NBER 数据库中的全部专利及识别出的专利睡美人文献的申请年度进行统计,并对各个年度及总体专利睡美人文献的占比进行计算得到表 3。由于 NBER 的数据在 2006 年之后便不再更新,本文所使用的 A. F. J. V. Raan 的识别方法对年份限制在至少 6 年的时间窗内进行计算,因此 2002 – 2006 年并没有出现专利睡美人文献的数据。从总体趋势来说,美国专利的申请量在不断增加,相应的专利睡美人文献的数量也在增加。但由于专利睡美人文献需要经过一段比较长的沉睡期之后,才能被发掘并获得重视,因此后期专利睡美人文献的数量减少。随着时间线的继续延长,沉睡专利被唤醒,这些年份的专利睡美人文献数量还将继续增加。

表 3 专利睡美人文献申请年度分布及占比情况

年份	NBER 专利量(件)	专利睡美人文献量(件)	专利睡美人文献所占比例(%)	年份	NBER 专利量(件)	专利睡美人文献量(件)	专利睡美人文献所占比例(%)
1967	215	0	0.00	1988	135 870	1 310	0.96
1968	256	0	0.00	1989	145 566	1 243	0.85
1969	428	2	0.47	1990	151 196	1 115	0.74
1970	779	1	0.13	1991	151 336	1 011	0.67
1971	1 678	6	0.36	1992	157 375	919	0.58
1972	3 191	7	0.22	1993	163 618	804	0.49
1973	12 617	28	0.22	1994	189 340	549	0.29
1974	557 06	131	0.24	1995	227 138	300	0.13
1975	93 638	219	0.23	1996	217 062	92	0.04
1976	95 667	246	0.26	1997	256 153	34	0.01
1977	97 182	263	0.27	1998	261 529	2	0.00
1978	98 343	295	0.30	1999	283 732	0	0.00
1979	99 652	379	0.38	2000	298 042	0	0.00
1980	100 134	457	0.46	2001	294 797	0	0.00
1981	96 984	507	0.52	2002	245 591	0	0.00
1982	98 559	557	0.57	2003	161 843	0	0.00
1983	93 591	591	0.63	2004	89 269	0	0.00
1984	102 809	737	0.72	2005	26 130	0	0.00
1985	109 216	914	0.84	2006	1 659	0	0.00
1986	114 620	1 090	0.95	合计	4 331 490	15 016	0.35
1987	123 471	1 207	0.98				

注:由于专利睡美人文献识别的时间窗不包含 2002 年 – 2006 年,因此合计一栏的数据也未将其列入计算

从专利睡美人文献所占比例来看,1981 年 – 1992 年,专利睡美人文献的年度占比都在 0.5% 以上,时间窗在 6 年以上的专利睡美人文献总体占比也在 0.35% 以上。随着时间的推移,这个比例将远远小于实际发生比例,说明有很大一批的专利在申请授权之初并没有被很好地利用,对于专利这类科技含量与创新性较高的科技文献来讲,这个比例不容小觑。部分年份每年有上千件的专利成为专利睡美人,在之后的 5 年甚至更长的时间内没有引起重视,给科学研究和技术发展造成了重大损失。

造成睡美人专利的原因是复杂的:与论文类似,专利的发明人与专利权人是否在相应技术研发领域具有权威,对专利的被引用情况产生的影响较大;睡美人专利所代表的技术能否与同时期技术发展阶段相匹配,也对专利能否被引用和取得成果转化起到重要作用,当睡美人专利在申请授权之初,其所代表的技术超前时,往往不能立即得到认可,当技术发展到相应阶段时,研究人员才能意识到该专利的价值;此外,该领域相关的社群氛围与态度能否接受新颖或独特的专利,也对专利的价值认可产生影响。专利本身就在一个复杂体系中运行,造成睡美人专利的原因相较于论文中的睡美人也更为多样。

4.1.2 专利睡美人文献睡眠长度 按照 A. F. J. V. Raan 的识别方法,专利睡美人文献的睡眠长度至少为 5 年,本研究统计结果见表 4。其中睡眠长度分布最为广泛的是 6 – 10 年,且专利数量均在 1 000 件以上,占专利睡美人文献总量的 59.89%,睡眠长度为 11 – 20 年的,占专利睡美人文献总量的 31.37%,睡眠长度为

20 年以上的，占专利睡美人文献总量的 4.14%。可见，睡眠长度在 6－20 年的专利睡美人文献占比最大。据此，在对某一领域的专利睡美人文献进行识别、预测与唤醒时，可将距离当前研究时间点前 6 年至前 20 年之间的专利文献为主要研究对象。时间越靠前的专利与当前研究在技术、方法上脱节越严重，其贡献与突破的可能性也将越小，睡眠专利被唤醒的成本与难度也将越大。

表 4　专利睡美人文献睡眠长度统计

睡眠长度(年)	专利量(件)	占比(%)
5	691	4.60
6	2 203	59.89
7	2 359	
8	1 853	
9	1 406	
10	1 172	
11	890	31.37
12	727	
13	618	
14	555	
15	425	
16	389	
17	332	
18	314	
19	255	
20	205	
21	172	4.14
22	111	
23	77	
24	89	
25	66	
26	46	
27	37	
28	17	
29	2	
30	3	
31	1	
32	1	
合计	15 016	100.00

4.1.3　专利睡美人文献唤醒强度　本研究对唤醒强度的设定为沉睡专利被唤醒后 4 年的年均引用量为 $C_w \geq 5$ 次。对识别出的专利睡美人文献的唤醒强度进行统计，结果见图 3。专利睡美人文献的唤醒强度在 5－10 次，即唤醒后年均引用量在 5－10 次的专利量是最普遍的，其次是唤醒强度为 10－20 次(不含 10 次)，唤醒强度为 20－30 次(不含 20 次)及 30－40 次(不含 30 次)的专利量相差不多且数量较少，唤醒强度在 40 次以上的专利占比最小。由此，绝大多数的专利睡美人文献的唤醒强度在 5－20 次之间，说明专利的唤醒是一个循序渐进的过程。我国学者李江认为睡美人文献唤醒强度越大，这种文献出现的概率越小，并且出现的概率大小与睡眠深度和睡眠长度之间没有直接关系[10]。从专利睡美人文献的角度来说，当专利本身价值很大，专利的技术贡献程度与技术发展相匹配时，更容易产生高唤醒强度。

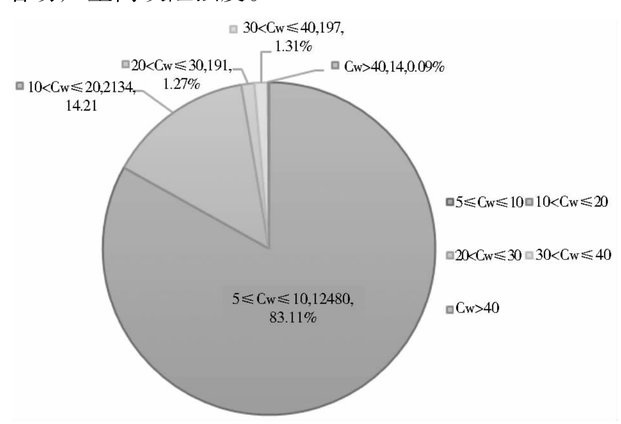


图 3　专利睡美人文献唤醒强度分布

4.1.4　专利睡美人文献国际分类号分布　国际专利分类号(IPC)是国际通用的专利文献分类和检索工具，一个完整的分类号由部、大类、小类和大组或小组的组合类号构成。专利有时不局限于某一固定领域，在专利申请与审查过程中，专利申请人与专利审查员将会给出与专利技术主题相对应的多个分类号。因此，同一件专利可以划分出两个或多个不同的分类。本文所识别出的专利睡美人文献的国际分类号记录共有 18 568 条，部别分类分析结果如图 4 所示：

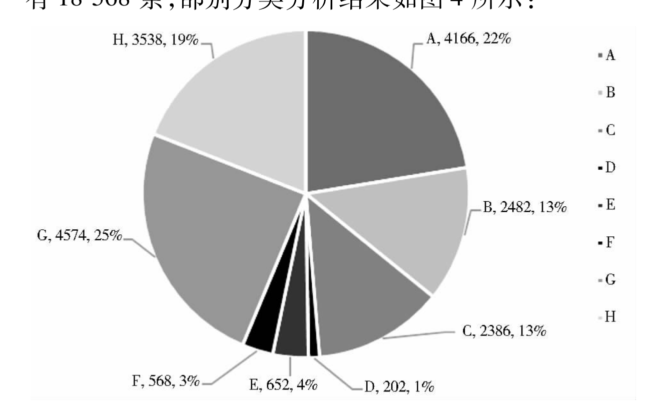


图 4　专利睡美人文献国际分类号部别分布

专利睡美人文献部别分类从多到少排列依次为：G(物理)、H(电学)、A(农业)、B(作业;运输)、C(化学;冶金)，占比在 10% －25%；E(固定建筑物)、F(机

械工程、照明、加热)、D(纺织、造纸)3 个部类的专利量均不足 5%，说明在 A、B、C、G、H 几个大部里产生的睡美人数量最多。

考虑到数据库中总体专利的部别分布，专利睡美人文献占相应部别专利总量的比例由大到小排序见表 5。据此分析排序靠前的 A、E、G、H 部更容易出现专利睡美人文献，而不一定是部别专利越多越容易产生专利睡美人文献，例如 B 类专利总量排名第二，而专利睡美人文献所占比例却排在倒数第二。

表 5 专利睡美人文献与整体专利部别分布类比

部别	专利总量(件)	专利睡美人文献量(件)	专利睡美人文献所占比例(%)
A	679 045	4 166	0.61
E	129 719	652	0.50
G	973 200	4 574	0.47
H	852 859	3 538	0.41
D	66 658	202	0.30
C	832 175	2 386	0.29
B	921 774	2 482	0.27
F	399 134	568	0.14
合计	4 854 564	18 568	0.38

注:表中专利总量是对部别进行降噪处理之后的数据，因此与数据库中的总体专利记录数存在差异

国际专利分类号的部别是从宏观上与产业相对应的划分方式，而大类的划分则相对较为详细，便于对专利所属的领域从更细粒度上进行判读。表 6 为大类分布前 20 的专利睡美人文献数量及所占比例情况。其中 A61(医学或兽医学;卫生学)、H01(基本电气元件)、H04(电通信技术)、G06(计算;推算;计数)和 G(测量;测试)所占比例较大，A61 的比例最大。与前面的结果类似，专利睡美人文献总量较大的部类，出现专利睡美人文献的概率不一定很大。

4.2 专利睡美人文献典型案例分析

本文从识别出的专利睡美人文献中选取典型个案进行分析。所选取的案例专利号为 4558333，检索时间为 2018 年 3 月 17 日，专利名称为“Liquid jet recording head”，即“液体喷射记录头”。申请日期为 1982 年 7 月，专利权人为日本佳能株式会社(Canon Kabushiki Kaisha)，国际专利分类号为 B41J 2/16 (20060101); G01D 015/18，对照《国际专利分类表第 8 版(2006.01)》[38]结果如下:

表 6 专利睡美人文献国际分类号专利量前 20 的大类分布

二级大类	含义	专利睡美人文献量(件)	占专利睡美人文献比例(%)	数据库中该类专利量(件)	专利睡美人文献占总体专利比例(%)
A61	医学或兽医学;卫生学	3 413	18.38	357 499	0.95
H01	基本电气元件	1 432	7.71	403 666	0.35
H04	电通信技术	1 332	7.17	233 493	0.57
G06	计算;推算;计数	1 311	7.06	244 851	0.54
G01	测量;测试	1 244	6.70	256 005	0.49
C08	有机高分子化合物;其制备或化学加工;以其为基料的组合物	600	3.23	169 398	0.35
C07	有机化学	579	3.12	273 819	0.21
G02	光学	562	3.03	97 282	0.58
E21	土层或岩石的钻进;采矿	525	2.83	35 989	1.46
G11	信息存储	477	2.57	104 553	0.46
G03	摄影术;电影术;利用了光波以外其他波的类似技术;电记录术;全息摄影术	400	2.15	125 723	0.32
B32	层状产品	396	2.13	80 352	0.49
C12	生物化学;啤酒;烈性酒;果汁酒;醋;微生物学;酶学;突变或遗传工程	367	1.98	126 378	0.29
B65	输送;包装;贮存;搬运薄的或细丝状材料	289	1.56	147 379	0.20
B01	一般的物理或化学的方法或装置	288	1.55	104 766	0.27
H02	发电、变电或配电	260	1.40	69 546	0.37
H03	基本电子电路	260	1.40	89 908	0.29
H05	其他类目不包含的电技术	243	1.31	54 166	0.45
B29	塑料的加工;一般处于塑性状态物质的加工	239	1.29	52 752	0.45
A63	运动;游戏;娱乐活动	211	1.14	57 247	0.37

B 作业;运输

B41 印刷;排版机;打字机;模印机

B41J 打字机;选择性印刷机构,即不用印版的印刷机构;排版错误的修正

B41J 2/00 以打印或标记工艺为特征而设计的打字机或选择性印刷机构

B41J 2/16…喷嘴的制造〔5〕

G 物理

G01 测量;测试

G01D 非专用于特定变量的测量;不包含在其他单独小类中的测量两个或多个变量的装置;计费设备;非专用于特定变量的传输或转换装置;未列入其他类目的测量或测试

G01D 15/00 非专用于特定变量的测量装置的记录器的组件

G01D 15/18…喷出记录材料的喷嘴

通过对专利号及专利内容的解析,发现该专利主要是对如打印机等印刷设备的喷嘴的制造、构造与使用材料进行的说明。该专利所申请授权的液体喷射记录头包括通过排放液体形成飞沫的排放孔;具有弯曲部分的液体通道,形成液体通道的能量作用部分,能量作用部分又包含了填充内部的滴状液体,以及用于产生待传输的液滴形成能量的能量产生元件,将能量传输到作用模块的填充液体中。液体通道壁表面的主要部分由硬化的光敏树脂构成。

通过对该专利的引文数据进行收集绘制引文曲线见图 5。该专利符合睡美人文献的典型特征,睡眠长度为 8 年,睡眠深度为 0.875,即该专利在睡眠期的年均被引仅为 0.875 次,在被唤醒之前为深睡状态。该专利于 1990 年被唤醒,被引用频次突然增加,并在随后的几年持续上升,专利的唤醒强度为 76.5,即被唤醒后 4 年内的年均被引次数为 76.5 次,可见该专利受到广泛关注,对技术革新和迭代产生了重要的作用。随着时间的延续和新技术的出现,引用量逐步削减,专利文献进入衰退期。

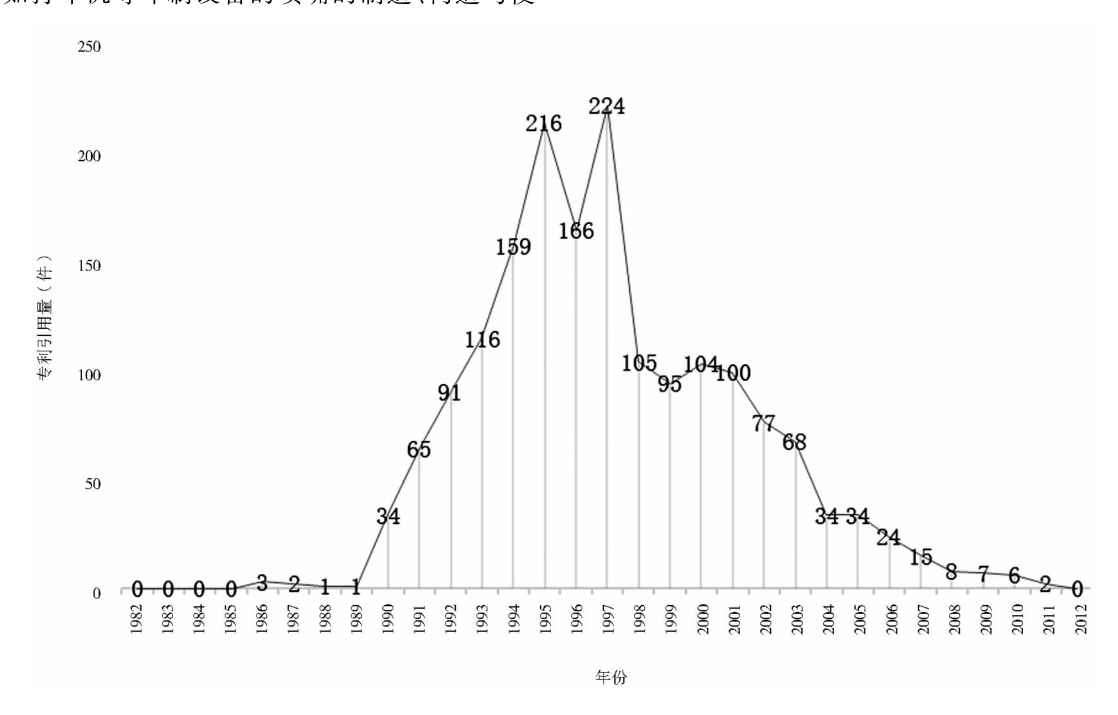


图 5　专利“Liquid jet recording head”的引文曲线

喷头、打印头是打印机的核心部件。1951 年西门子最早申请喷墨打印专利,但在很长一段时间内并没有实现商业化,直到 1977 年西门子才推出第一款商用喷墨打印机。同时期,日本佳能公司也在进行相关研究,于 1982 年开始实施第二个“优良企业”构想计划,并推出世界首创暗合方式小型复印机“PC10/PC - 20”等一系列成果[39],可见该专利就是佳能公司在该时期的研究成果之一。在该专利的睡眠期,即 1982 年 - 1989 年间,引用该专利的主要是其竞争对手同时也是 1985 年与之结为战略联盟的合作伙伴——惠普公司(见表 7),可见惠普对佳能的关注程度之高。此外,佳能(自引)、柯达和光谱公司分别对该专利施引一次,说明当时在液体打印喷头方向这些公司参与了研发和布局并关注和借鉴了佳能的成果。

表 7　专利“Liquid jet recording head”沉睡期前向引用专利信息

专利号	专利名称	申请时间（年）	专利权人	国际专利分类号
4680859	Thermal ink jet print head method of manufacture	1986	Hewlett-Packard Company（Palo Alto，CA）	B41J 2/14；G01D 015/16；H05B 003/00；B05D 005/12；G03C 005/00
4694308	Barrier layer and orifice plate for thermal ink jet print-head assembly	1986	Hewlett-Packard Company（Palo Alto，CA）	B41J 2/16；B41J 2/14；C25D 1/00；C25D 1/02；G01D 015/18
4716423	Barrier layer and orifice plate for thermal ink jet print head assembly and method of manufacture	1986	Hewlett-Packard Company（Palo Alto，CA）	B41J 2/14；B41J 2/16；C25D 1/00；C25D 1/02；G01D 015/18；C25D 001/02；G03C 005/00
4794410	Barrier structure for thermal ink-jet printheads	1987	Hewlett-Packard Company（Palo Alto，CA）	B41J 2/20；B41J 2/17；B41J 2/14；G01D 015/18
4835554	Ink jet array	1987	Spectra，Inc.（Hanover，NH）	B41J 2/515；B41J 2/145；B41J 2/155；B41J 2/175；B41J
4905017	Laminated liquid-jetting head capable of recording in a plurality of colors，a method of producing the head and an apparatus having the head	1988	Canon Kabushiki Kaisha（Tokyo，JP）	B41J 2/14；B05B 1/14；G01D 015/16；B41J 003/04
4942408	Bubble ink jet print head and cartridge construction and fabrication method	1989	Eastman Kodak Company（Rochester，NY）	B41J 2/14；B41V 002/05；B41V 002/175

该专利于 1990 年被唤醒，绝大部分的引用来自于佳能公司的自引用，其次是惠普公司，但其占比相较于自引率微乎其微。此外日本最大的复印机制造企业理光微公司（Ricoh Co.，Ltd.）、全球最大现代化办公设备制造商施乐公司（Xerox Corporation）、精工爱普生公司（Seiko Epson Corporation）、谷歌公司（Google Inc.）等也对该专利进行过少量引用。专利的自引率高说明该专利的专利权人在本技术领域的研发和应用及成果转化上具有高参与度以及很强的自我继承能力[40]，佳能的此项专利是佳能在打印机领域早期的重要研发成果，为后续企业发展乃至行业技术发展都做出了相当大的贡献。随着打印机技术的发展与成熟，打印头的研发投入与研发方向的变化，是专利近年来被引用量降低的主要原因。

总之，该专利在申请并得到授权之初并未受到广泛关注，说明在当时技术的发展过程中以及佳能的研发进程中还未能将专利技术转化，该专利具有一定的超前性。而随着佳能对打印头的研究继续深入，该专利的价值得以凸显。该专利从沉睡到唤醒，专利价值得以发挥，而当技术发展成熟之后，专利的参考价值也将逐步降低，最终表现为引用量的减少。

5　总结与展望

专利信息挖掘与成果转化是当前面临的重要问题，文章借鉴科学睡美人文献的相关研究思路与分析方法，以美国 NBER 专利数据计划的相关数据和美国专利商标局的专利数据为基础，探究专利文献中的睡美人现象，并从整体上进行专利睡美人文献的识别与特征分析：

（1）本文所识别的专利睡美人文献所占比例略大于普通论文识别的睡美人文献的比例，在一定程度上说明科学论文和技术文献存在差异性，专利文献受专利制度和可获得性等因素影响，相较于科学论文更容易产生沉睡现象。

（2）专利睡美人文献的睡眠长度主要集中在 6 – 20 年之间。本文使用的引证数据截止到 2006 年，所识别出的专利睡美人文献主要分布在 1981 年 – 1992 年。睡眠长度过长的专利往往并不能很好地为前沿科技发展提供有效支撑，但有效时间内的专利回溯与信息挖掘产生突破性发现的可能性较大。对此类专利进行研究在一定程度上有助于加快技术发展进程，推动科技快速发展，并减少科研浪费，提高科研效率。本研究发现，专利睡美人文献的唤醒强度越大，专利数量越少。当专利本身价值很大，专利的技术贡献程度与技术发展相匹配时，更容易产生高唤醒强度。此外，对国际分类号分布研究发现，专利睡美人文献总量较大的部类，出现睡美人专利的概率不一定很大。

（3）通过对典型案例的分析发现，专利文献的沉睡与唤醒可能与企业的发展战略、市场竞争活动、技术发展以及产业结构变化等具有相关性。此外，受专利制度的影响，专利文献的引用动机和与科学论文存在诸多不同[41]，当竞争对手引用某一专利，在一定程度上是为了规避该专利技术。因此，不能简单认为高被引专利文献一定就是有价值的专利，或称之为“延迟认

可”,而批评、反驳的态度也不代表被引专利文献存在“错误”。事实上,引证关系本身就代表了对被引用一方的关注,体现了知识的链接和技术的迭代,具有现实意义。

本文通过对专利睡美人文献的识别和特征分析,为后续识别与唤醒睡美人专利提供了研究基础。着重关注睡美人专利出现可能性较大的领域,可有效地识别、关注并及时唤醒沉睡专利,助力新技术创新。本文的局限性在于没有排除专利的自引,所使用的引证数据并未更新到当前时间点,研究结果可能会随着时间的推移与技术的发展在特征上产生差异。但从历史数据的信息挖掘方面来讲,已产生的引证数据并不会受较大影响。本文的目的是尝试性识别并揭示专利中的睡美人现象,未来研究将关注睡美人专利与睡美人论文存在的差异,区分专利文献的引用动机和引证效果,在识别方法上结合专利自身特征予以改进和优化,以期为睡美人专利的识别和唤醒做出新的贡献,为专利信息分析的理论与实践提供新的角度。

参考文献:

[1] 杜建.“睡美人”文献的识别方法与唤醒机制研究[D]. 南京:南京大学,2017.
[2] JAFFE A B, TRAJTENBERG M, FOGARTY M S. Knowledge spillovers and patent citations: evidence from a survey of inventors[J]. American economic review, 2000, 90(2):215－218.
[3] CHEN L. Do patent citations indicate knowledge linkage? The evidence from text similarities between patents and their citations[J]. Journal of informetrics, 2017, 11(1):63－79.
[4] AMIN M, MABE M. Impact factors: use and abuse[J]. International journal of environment science and technology, 2000, 1(1): 1－6.
[5] EGGHE L, RAO I K R. Citation age data and the obsolescence function: fits and explanations[J]. Information processing & management, 1992, 28(2):201－217.
[6] BARBER B. Resistance by scientists to scientific discovery[J]. Science, 1961, 134(3479): 596－602.
[7] STENT G S. Prematurity and uniqueness in scientific discovery[J]. Scientific American, 1972, 227(6): 84－93.
[8] COLE S. Professional standing and the reception of scientific discoveries[J]. American journal of sociology, 1970, 76(2): 286－306.
[9] VAN RAAN A F J. Sleeping Beauties in science[J]. Scientometrics, 2004, 59(3):467－472.
[10] 李江. 科学中的“睡美人”与“昙花一现”现象评述[J]. 大学图书馆学报,2016,34(3):38－43.
[11] LI J, YE F Y. The phenomenon of all-elements-sleeping-beauties in scientific literature[J]. Scientometrics, 2012, 92(3):795－799.
[12] COSTAS R, VAN LEEUWEN T N, VAN RAAN A F J. Is scientific literature subject to a ‘sell-by-date’? A general methodology to analyze the ‘durability’ of scientific documents[J]. Journal of the Association for Information Science and Technology, 2010, 61(2): 329－339.
[13] KE Q, FERRARA E, RADICCHI F, et al. Defining and identifying sleeping beauties in science[J]. Proceedings of the National Ccademy of Sciences, 2015, 112(24): 7426－7431.
[14] 李江,姜明利,李玥婷. 引文曲线的分析框架研究——以诺贝尔奖得主的引文曲线为例[J]. 中国图书馆学报,2014,40(2):41－49.
[15] WANG J, MA F, CHEN M, et al. Why and how can “sleeping beauties” be awakened? [J]. The electronic library, 2012, 30(1): 5－18.
[16] 杜建,武夷山. 睡美人与王子文献的识别方法研究[J]. 图书情报工作,2015,59(19):84－92.
[17] GILBERT R J, NEWBERY D M G. Preemptive patenting and the persistence of monopoly[J]. The American economic review, 1982,72(3): 514－526.
[18] LEUNG C M, KWOK Y K. Real options game analysis of sleeping patents[J]. Decisions in economics and finance, 2011, 34(1): 41－65.
[19] PALOMERAS N. Sleeping patents: any reason to wake up[R]. Barcelona: IESE Business School, 2003. NO. d/506.
[20] WEEDS H. Sleeping patents and compulsory licensing: an options analysis[R]. Coventry: University of Warwick, 1999. NO. 577.
[21] 唐要家,孙路. 专利转化中的“专利沉睡”及其治理分析[J]. 中国软科学,2006(8):73－78.
[22] 袁晓东. 沉睡专利形成机理及其防治[J]. 科研管理,2009,30(4):168－174.
[23] 张宇青. 我国“专利沉睡”之困与治理研究[J]. 科学管理研究,2013,31(4):49－52.
[24] SOININEN A H. Open standards and the problem with submarine patents[C]// The 4th conference on standardization and innovation in information technology. Geneva: IEEE, 2005: 218－231.
[25] 孙兆刚. 潜水艇专利的规避对策研究[J]. 科技管理研究,2012,32(9):167－170.
[26] 唐雯,刘超. 专利条约下“潜水艇专利”策略及其合法性分析[J]. 行政与法,2012(9):104－108.
[27] 雷敏敏. 论美国专利侵权中的懈怠抗辩及其借鉴[D]. 北京:北京化工大学,2011.
[28] 栾春娟,侯海燕,王贤文. 重大发明创造的技术多元化特征更明显吗? [J]. 科学学与科学技术管理,2014,35(4):66－73.
[29] SEIDEL A H. Citation system for patent office[J]. Journal of the Patent Office Society, 1949, 31(5): 554.
[30] 杨中楷. 专利计量与专利制度[M]. 大连:大连理工大学出版社, 2008:21－22.
[31] NARIN F. Patent bibliometrics[J]. Scientometrics, 1994, 30

(1): 147 - 155.

[32] 文庭孝. 专利信息计量学[M]. 北京:科学出版社, 2017:8.

[33] ARMANDO A, PLAZA L M. The transfer of knowledge from the Spanish public R&D system to the productive sectors in the field of biotechnology[J]. Scientometrics, 2004, 59(1): 3 - 14.

[34] 高继平,丁堃. 专利计量指标研究述评[J]. 图书情报工作, 2011,55(20):40 - 43.

[35] 马廷灿,李桂菊,姜山,等. 专利质量评价指标及其在专利计量中的应用[J]. 图书情报工作,2012,56(24):89 - 95,59.

[36] 张国富,韩宁. "沉睡专利"之成因与激活路径探讨[J]. 学习论坛,2013,29(8):70 - 73.

[37] HALL B H, JAFFE A B, TRAJTENBERG M. The NBER patent citation data file: lessons, insights and methodological tools[R]. Cambridge: National Bureau of Economic Research, 2001.

[38] 中华人民共和国国家知识产权局. 国际专利分类表第 8 版 (2006.01) [EB/OL]. [2017 - 12 - 28]. http://www.sipo.gov.cn/wxfw/zlwxxxggfw/zsyd/bzyfl/gjzlfl/201406/t20140630_973348.html.

[39] 佳能中国. 集团历史[EB/OL]. [2018 - 01 - 01]. http://www.canon.com.cn/about/history/overview.

[40] 赵亚娟. 专利引用分析方法与应用[J]. 图书情报工作,2009,53(6):11 - 15.

[41] 李睿,孟连生. 论专利引用行为与期刊论文引用行为在揭示知识关联方面的差异[J]. 情报学报,2010(3):474 - 478.

作者贡献说明:

李贺:研究方向的确定,研究方案指导,论文修改与完善;

袁翠敏:研究方案设计,数据收集分析与论文撰写;

解梦凡:数据分析处理。

Analysis and Study on Phenomenon of Sleeping Beauty in Patent Literature

Li He Yuan Cuimin Xie Mengfan

School of Management, Jilin University, Changchun 130022

Abstract: [**Purpose/significance**] The sleeping beauty literature is a description of the phenomenon of delayed recognition in scientific literature, and delayed discovery or delayed attention also exists in technical literature. After reviewing the research of sleeping beauty in the literature, sleeping patent phenomenon and patent citation analysis, this article introduces this concept into patent information analysis and explores the phenomenon of sleeping beauty in patent documents. [**Method/process**] Based on the database of patents and patent citations published by the United States Patent and Trademark Office and the National Bureau of Economic Research, the paper uses classical identification method of sleeping beauty to identify the sleeping beauty patents, studies characteristics of this kind of patents, and conducts a case study by selecting a typical case. [**Result/conclusion**] The results prove that sleeping beauty phenomenon exists in patent literature while with unique characteristics. And this research lays a foundation for the subsequent identification and awakening of sleeping beauty patents, and then provides solutions for early detection and utilization of such valuable patent documents, facilitates the flow of knowledge and the iteration of technology, promotes scientific discovery and scientific research efficiency, and accelerates research process.

Keywords: delayed recognition sleeping beauty literature patent citation analysis sleeping beauty patent